$$\vec{Y} = X\beta + \epsilon$$

$$\downarrow$$

Lec 10 MATH 393

$$\vec{Y} = X\vec{\beta} + \vec{\mathcal{E}}$$ Assume $\vec{\mathcal{E}} \sim N_n(\vec{0}_n, \sigma^2 I_n)$

$\Rightarrow \mathcal{E}_1, \ldots, \mathcal{E}_n \overset{iid}{\sim} N(0,1)$

OGP for misspecification error + ignorance error

$\vec{Y}$: random through $\vec{\mathcal{E}}$; X: fixed constants; $\vec{\beta}$: unknown parameter

$$\Rightarrow \vec{B} = (X^\top X)^{-1} X^\top \vec{Y} \sim N_{p+1}\left(\vec{\beta}, \sigma^2 (X^\top X)^{-1}\right) \Rightarrow B_j \sim N\left(\beta_j, \sigma^2 (X^\top X)^{-1}_{jj}\right)$$

$$\Rightarrow \frac{B_j - \beta_j}{\sigma\sqrt{(X^\top X)^{-1}_{jj}}} \sim N(0,1) \quad \forall j$$

But σ is unknown!

$$\Rightarrow \hat{\vec{Y}} = X\vec{B}$$

$$\Rightarrow \vec{Y} = X\vec{B} + \vec{E} = \hat{\vec{Y}} + \vec{E}$$

$\vec{E}$: OGP for $\vec{e}$, the residuals in the regression

Left over from lec 9:

$$\Rightarrow \vec{E} = \vec{Y} - X\vec{B} \underset{\text{geometrically}}{=} (I - H)\vec{Y} = (I-H)(X\vec{\beta} + \vec{\mathcal{E}}) = \underbrace{(I-H)X\vec{\beta}}_{\vec{0}_{(p+1)\times(p+1)}} + (I-H)\vec{\mathcal{E}}$$

$$= (I-H)\vec{\mathcal{E}} \sim N_n\left(\vec{0}_n, (I-H)\sigma^2 I_n (I-H)^\top\right) = N_n\left(\vec{0}, \sigma^2(I-H)\right)$$

$$\Rightarrow E_i \sim N(0, \sigma^2(1 - H_{ii}))$$

From MATH 390: $\vec{U} \sim N_n(\vec{0}_n, \Sigma)$, $\vec{\mu} \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n} \Rightarrow \vec{\mu} + A\vec{U} \sim N_m(\vec{\mu}, A\Sigma A^\top)$

$$\hat{\vec{Y}} = X\vec{B} = H\vec{Y} = H(X\vec{\beta} + \vec{\mathcal{E}}) = HX\vec{\beta} + H\vec{\mathcal{E}} = X\vec{\beta} + H\vec{\mathcal{E}} \sim N_n(X\vec{\beta}, H\sigma^2 I_n H) = N_n(X\vec{\beta}, \sigma^2 H) \Rightarrow \hat{Y}_i \sim N(\vec{x}_i \cdot \vec{\beta}, \ldots$$

In 390 we showed that if $\vec{Z} \sim N_n(\vec{0}, \sigma^2 I_n)$

$$\Rightarrow \vec{Z}^\top \vec{Z} \sim \chi^2_n$$

From last class...

$$\vec{Z}^\top \vec{Z} = \frac{1}{\sigma^2} \vec{\mathcal{E}}^\top \underbrace{H}_{B_1} \vec{\mathcal{E}} + \frac{1}{\sigma^2} \vec{\mathcal{E}}^\top \underbrace{(I-H)}_{B_2} \vec{\mathcal{E}}$$

Since $B_1 + B_2 = I_n$ and $\text{rank}(B_1) + \text{rank}(B_2) = (p+1) + (n-(p+1)) = n$, we can use Cochran's thm

$$\Rightarrow \frac{1}{\sigma^2} \vec{\mathcal{E}}^\top H \vec{\mathcal{E}} \sim \chi^2_{p+1} \text{ indep of } \frac{1}{\sigma^2} \vec{\mathcal{E}}^\top (I-H) \vec{\mathcal{E}} \sim \chi^2_{n-(p+1)}$$

last class ↓ measures ||| error

$$\frac{1}{\sigma^2}\vec{\varepsilon}^\top(I-H)\vec{\varepsilon} = \dots\dots = \frac{1}{\sigma^2}\|\vec{E}\|^2 \sim \chi^2_{n-(p+1)}$$

$$\frac{1}{\sigma^2}\vec{\varepsilon}^\top H\vec{\varepsilon} = \frac{1}{\sigma^2}\vec{\varepsilon}^\top HH\vec{\varepsilon} = \frac{1}{\sigma^2}\vec{\varepsilon}^\top H^\top H\vec{\varepsilon} = \frac{1}{\sigma^2}(H\vec{\varepsilon})^\top(H\vec{\varepsilon}) = \frac{1}{\sigma^2}\|H\vec{\varepsilon}\|^2$$

$$= \frac{1}{\sigma^2}\|H(\vec{Y}-X\vec{\beta})\|^2 = \frac{1}{\sigma^2}\|H\vec{Y}-HX\vec{\beta}\|^2 = \frac{1}{\sigma^2}\|\hat{\vec{Y}}-X\vec{\beta}\|^2$$

$$= \frac{1}{\sigma^2}\|\underbrace{X\vec{B}}_{\hat{Y}}-\underbrace{X\vec{\beta}}_{E[Y]}\|^2 = \frac{1}{\sigma^2}\|X(\vec{B}-\vec{\beta})\|^2 \sim \chi^2_{p+1}$$

measures estimation error

As $n\to\infty$ $\vec{E}\to\vec{\varepsilon}$, $\vec{B}\to\vec{\beta} \Rightarrow \|X(\vec{B}-\vec{\beta})\|^2\to 0$

misspecification + ignorance error. estimation error disappears

$\Rightarrow \vec{E}$ and $\vec{B}$ are independent $\Rightarrow \|\vec{E}\|^2$ and B_j are independent

Consider the following estimator

from last class

$$\frac{\overbrace{\dfrac{B_j-\beta_j}{\sigma\sqrt{(X^\top X)^{-1}_{jj}}}}^{\sim N(0,1)}}{\sqrt{\dfrac{\overbrace{\frac{1}{\sigma^2}\|\vec{E}\|^2}^{\sim\chi^2_{n-(p+1)}}}{n-(p+1)}}} \sim T_{n-(p+1)}$$

indep. we just proved this

Student's T distr

$$= \frac{\frac{1}{\sigma}\dfrac{B_j-\beta_j}{\sqrt{(X^\top X)^{-1}_{jj}}}}{\frac{1}{\sigma}\sqrt{\dfrac{\|E\|^2}{n-(p+1)}}} \sim T_{n-(p+1)}$$

the estimated test statistic is then

$$\frac{\dfrac{b_j-\beta_j}{\sqrt{(X^\top X)^{-1}_{jj}}}}{} \qquad \frac{b_j-\beta_j}{}$$

Assume level of test is α:

Usually the inference is for $H_a: \beta_j \neq 0 \Leftrightarrow H_0: \beta_j = 0$ why?

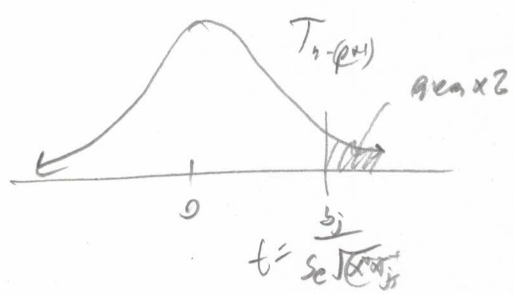
Because if $\beta_j = 0$, the j^{th} feature doesn't affect the response (linearly)

(which usually means the overall relationship is not too interesting).

In this case,

$$\frac{b_j}{s_e\sqrt{(X^TX)^{-1}_{jj}}} \in \left[\pm t_{1-\frac{\alpha}{2},\, n-(p+1)}\right] \xrightarrow{\text{inverting the test}} CI_{\beta_j,1-\alpha} = \left[b_j \pm t_{1-\frac{\alpha}{2},\, n-(p+1)} \cdot s_e\sqrt{(X^TX)^{-1}_{jj}}\right]$$

How to get pval?



pval =

$$2P\left(|T_{n-(p+1)}| > \left|\frac{b_j}{s_e\sqrt{(X^TX)^{-1}_{jj}}}\right|\right)$$

Consider a new $\vec{x}_*$ and you want to predict its expected response $h^*(\vec{x}_*) = \vec{x}_* \vec{\beta}$. This is the average of many predictions of $\vec{x}_*$ since the noise mean is zero.

Let $\mu_* := \vec{x}_* \vec{\beta}$

$$\hat{Y}_* = \vec{x}_* \vec{B} \Rightarrow \hat{Y}_* \sim N(\underbrace{\vec{x}_* \vec{\beta}}_{\mu_*}, \sigma^2 \vec{x}_* (X^T X)^{-1} \vec{x}_*^T)$$

$$\vec{B} \sim N(\vec{\beta}, \sigma^2 (X^T X)^{-1})$$

$$\Downarrow$$

$$\frac{\hat{Y}_* - \mu_*}{\sigma \sqrt{\vec{x}_* (X^T X)^{-1} \vec{x}_*^T}} \sim N(0,1)$$

Again, σ is unknown so consider

since $\hat{Y}_*$ is a function of $\vec{B}$.

$$\frac{\overbrace{\dfrac{\hat{Y}_* - \mu_*}{\sigma \sqrt{\vec{x}_* (X^T X)^{-1} \vec{x}_*^T}}}^{N(0,1)}}{\underbrace{\sqrt{\dfrac{\frac{1}{\sigma^2} \|\vec{E}\|^2}{n-(p+1)}}}_{\chi^2_{n-(p+1)}}} \sim T_{n-(p+1)} \quad \text{indep}$$

$$\|$$

$$\frac{\dfrac{\hat{Y}_* - \mu_*}{\sqrt{\vec{x}_* (X^T X)^{-1} \vec{x}_*^T}}}{\sqrt{\dfrac{\|\vec{E}\|^2}{n-(p+1)}}} \sim T_{n-(p+1)}$$

the test statistic for $H_a: \mu_* \neq \theta$ is

$$\text{if } \frac{\hat{Y}_* - \theta}{s_e \sqrt{\vec{x}_* (X^T X)^{-1} \vec{x}_*^T}} \in \left[\pm t_{1-\frac{\alpha}{2}, n-(p+1)}\right] \Rightarrow \text{Retain}$$

which can be used to test. Inverting the test yields

$$CI_{\mu_*, 1-\alpha} = \left[\hat{Y}_* \pm t_{1-\frac{\alpha}{2}, n-(p+1)} \cdot s_e \sqrt{\vec{x}_* (X^T X)^{-1} \vec{x}_*^T}\right]$$